农业复杂环境下尺度自适应小目标识别算法——以蜜蜂为研究对象

郭秀明，诸叶平，李世娟，张　杰，吕纯阳，刘升平*

（中国农业科学院农业信息研究所/农业农村部农业信息服务技术重点实验室，北京 100081）

摘　要： 农业生产环境中的目标识别对象常具有分布密集、体积小、密度大的特点，加之农田环境光照多变、背景复杂，导致已有目标检测模型无法取得令人满意的效果。本研究以提高小目标的识别性能为目标，以蜜蜂识别为例，提出了一种农业复杂环境下尺度自适应小目标识别算法。算法克服了复杂多变的背景环境的影响及目标体积较小导致的特征提取困难，实现目标尺度无关的小目标识别。首先将原图拆分为一些较小尺寸的子图以提高目标尺度，将已标注的目标分配到拆分后的子图中，形成新的数据集，然后采用迁移学习的方法重新训练并生成新的目标识别模型。在模型的使用中，为使子图识别结果能正常还原，拆分的子图之间需具有一定的重叠率。收集所有子图的目标识别结果，采用非极大抑制（Non-Maximum Suppression，NMS）去除由于模型本身产生的冗余框，提出一种交小比非极大抑制（Intersection over Small NMS，IOS-NMS）进一步去除子图重叠区域中的冗余框。在子图像素尺寸分别为300×300、500×500和700×700，子图重叠率分别为0.2和0.05的情况下进行验证试验，结果表明：采用SSD（Single Shot MultiBox Detector）作为框架中的目标检测模型，新提出的尺度自适应算法的召回率和精度普遍高于SSD模型，最高分别提高了3.8%和2.6%，较原尺度的YOLOv3模型也有一定的提升。为进一步验证算法在复杂背景中小目标识别的优越性，从网上爬取了不同尺度、不同场景的农田复杂环境下的蜜蜂图像，并采用本算法和SSD模型进行了对比测试，结果表明：本算法能提高目标识别性能，具有较强的尺度适应性和泛化性。由于本算法对于单张图像需要多次向前推理，时效性不高，不适用于边缘计算。

关键词： 目标检测；机器视觉；小目标；农业环境；蜜蜂；SSD；YOLOv3

中图分类号： S24;TP391.41　　**文献标志码：** A　　**文章编号：** SA202203003

引用格式：郭秀明，诸叶平，李世娟，张杰，吕纯阳，刘升平. 农业复杂环境下尺度自适应小目标识别算法——以蜜蜂为研究对象[J]. 智慧农业(中英文), 2022, 4(1): 140-149.

GUO Xiuming, ZHU Yeping, LI Shijuan, ZHANG Jie, LYU Chunyang, LIU Shengping. Scale adaptive small objects detection method in complex agricultural environment: Taking bees as research object[J]. Smart Agriculture, 2022, 4(1): 140-149. (in Chinese with English abstract)

1 引 言

随着卷积神经网络及深度学习技术的发展[1]，基于机器视觉的目标检测受到了广泛关注，已取得了突破性进展[2,3]。农业中存在着许多目标物识别和计数的场景，用机器视觉技术对

收稿日期：2021-12-31

基金项目：中央级公益性科研院所基本科研业务费专项（2021JKY038）；河北省重点研发计划项目子课题（19227407D）；国家蜂产业技术体系专项（CARS-44）

作者简介：郭秀明（1981—），女，博士，副研究员，研究方向为农业智能感知。E-mail：guoxiuming@caas.cn。

*通信作者：刘升平（1981—），男，博士，研究员，研究方向为农产品质量安全控制、作物模拟模型技术等。E-mail：liusheng-ping@caas.cn。

农业中的目标物进行智能识别和计数能提高农业的智能化和现代化水平。农业生产环境多为室外环境，光线多变，背景复杂，且农业生产环境中的目标物大多具有体积小、密度高的特点。复杂背景环境下小目标的识别和检测是农业生产环境中常见的应用场景，如农业遥感图像中小目标、果树上的果实、蜂巢内的蜜蜂等，面向农业特定应用场景的要求，研究特定需求的算法模型以获得其在某一侧重指标的优越性能是未来几年内农业智能识别领域的研究趋势。

由于小目标的有效像素少、尺度小，缺乏特征表达能力，其检测一直是目标检测中的难点。已有众多研究者从不同角度设计优化检测模型[4]以提高小目标的检测性能。有的优化和改进主干网络结构[5-11]以提取更丰富的特征，有的优化锚框[12-17]以提高目标的定位精度，有的优化损失函数[18-20]以提高模型的训练效率和模型性能，这些改进方法能一定程度地提高对小目标的识别性能。然而，小目标的像素少且尺度小是造成其识别性能较差的根本原因。增加小目标的有效像素数以及增加其尺度是改善其识别性能的主要途径。同时，由于网络输出层包含全连接层，当前的网络模型大多都会对输入图像归一化至标准尺度，如快速区域卷积神经网络（Faster-Regions with Convolutional Neural Network，Faster-RCNN）[21]和SSD（Single Shot MultiBox Detector）[22]。归一化处理会导致图像尺寸进一步缩小，目标的有效像素进一步减少，增加了小目标识别的难度。提高输入图像中的小目标的尺度，增加小目标的像素个数，有利于充分提取其特征并提高其识别性能。已有的方法大多通过优化模型提高小目标的识别性能，较少从提高小目标的尺度（目标像素数与整体图像像素数的比值）出发进行研究。本研究着眼于影响小目标识别性能不高的本质原因——有效像素少且尺度小，通过图像拆分的方法简单有效地提高小目标的尺度，以期提高其识别性能。

蜜蜂体积小且在图像中尺度小，且蜜蜂常会聚集成簇分布，是农业中典型的小目标识别计数场景。本研究以蜂巢口的蜜蜂为例，提出了一种与输入图像尺寸和目标尺度无关的基于图像拆分的小目标识别算法。首先将原始输入图像拆分为多个子图，相邻子图之间设置有重叠区域，将多个子图作为模型的输入图像，将子图的输出结果汇集，然后采用二阶段非极大抑制（Non-Maximum Suppression，NMS）方法去除由于模型本身及子图重叠产生的冗余框。为评估算法性能，利用本算法和SSD及YOLOv3（You Only Look Once）模型进行了验证试验；从网上爬取了多种尺度和背景下的蜜蜂图像，将本算法和SSD模型进行了对比测试，评估了算法尺度适应能力及泛化性。

2 尺度自适应小目标识别算法

2.1 算法框架介绍

基于深度学习的目标检测算法主要分为前处理、特征提取和后处理3个部分（图1）。传统算法直接将整体图像作为网络模型的输入，为增强困难小目标的识别性能，本算法将输入图像拆分为若干子图，提高小目标的尺度，增加其像素数量。

后处理主要使用NMS去除卷积神经网络输出的冗余候选框，找到最佳的目标位置，提高检测的准确率。NMS是基于深度学习的目标检测中非常重要的一步，最早提出的NMS[23]将所有的候选框按得分值从高到低排序，选取得分值最高的候选框，删除所有与其重叠率超出设定阈值的候选框，对未删除的候选框选取得分值最高的继续此操作。此处的重叠率取值为相邻两个框的交并比（Intersection over Union，IOU），即两个框的交集面积与其并集面积的比值。针对不同的应用场景中NMS存在的问题，分别对其进行改进提出了Soft-NMS[24]、A-NMS[25]等多种非极大值抑制算法。本研究的算法不仅有深度学习网络模型产生的冗余框，还有图像的区域重叠造成

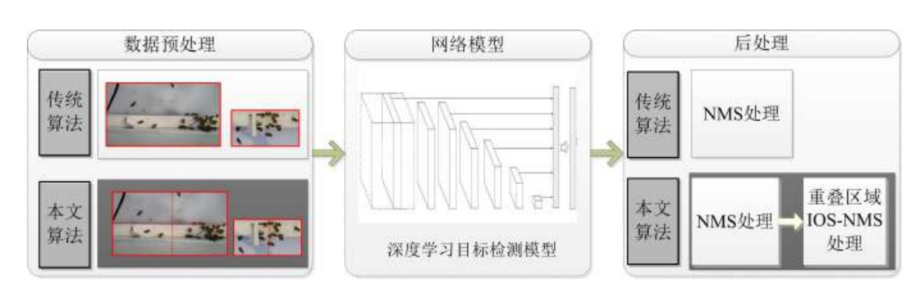


图1 尺度自适应小目标识别算法和传统基于深度学习的目标检测算法框架对比图

Fig. 1 Framework comparison between scale adaptive small objects detection algorithm and the traditional object detection algorithms based on deep learning

的检测冗余，针对后者，提出了一种交小比非极大抑制（Intersection over Small NMS，IOS-NMS）方法以进一步准确地定位目标位置。本算法和传统基于深度学习的目标检测算法框架对比图见图1。

2.2 新数据集生成方法

数据于2020年在中国农业科学院农业信息研究所采集，拍摄时间在蜜蜂较活跃的6月份。蜂巢口是蜂巢与外界的交界处，光线不受遮挡且蜜蜂较为活跃，将摄像头置于蜂巢口正上方，摄像头分辨率为1280×720像素，采集时间从早上8点持续到下午6点，拍摄间隔为45 s。涉及了早、中、晚多个时间段和多种天气状况，共获取有效图像2613张。采用手工标注对原图进行蜜蜂标注，如图2（a）所示。以目标识别模型SSD为例，原图输入模型后首先对其进行归一化为300×300像素的图像（图2（b）），归一化后蜜蜂尺度均值为14×18像素，而SSD模型中面积最小锚框像素大小为30×30，即使最小锚框也是蜜蜂像素尺度的3.57倍，造成蜜蜂回归位置不准确，识别性能下降。

为增加小目标的有效像素，提高其尺度，使用网格划分的方法将原图拆分，拆分子图的个数和子图的尺寸、相邻子图的重叠率有关。新生成的子图集合产生新的数据集用于模型训练（图3）。子图尺寸与模型的归一化输入尺寸、目标的尺度及原图的分辨率有关。为避免正负样本比例不均衡，提高数据输入的有效性，加快模型



（a）采集及标注的原图　（b）SSD模型归一化图像

图2 蜜蜂原图和归一化后图像及最小锚框示意图

Fig. 2 Labeled images before and after scale normalization and the minimum anchor box

的训练效率，移除没有目标物的子图，将含有目标的子图加入到新数据集中。由于蜜蜂标注是在原图像上进行的，新数据集生成中需要针对子图对标注信息进行重新计算，算法流程如下所示。

（1）原数据集设为A，对于任何一个原图∀a∈A；

（2）设a的宽度为w，高度为h，a中的目标物为集合O，包含有目标的位置信息和类别信息。设定目标子图的宽度为zw，高度为zh；

（3）对于原图，水平方向以zw为间隔，垂直方向以zh为间隔，将其划分为$\lceil w/zw \rceil \times \lceil h/zh \rceil$个子图，边缘处的子图剩余部分用纯色填充；

（4）对∀o∈O，对其进行重分配和坐标的重新计算，其中o是O中的一个元素；

（5）从中提取存在目标的子图加入新建数据集B中。

子图拆分时，需将原目标进行重新分配并调整其在子图中的坐标，目标重分配过程如图4所示

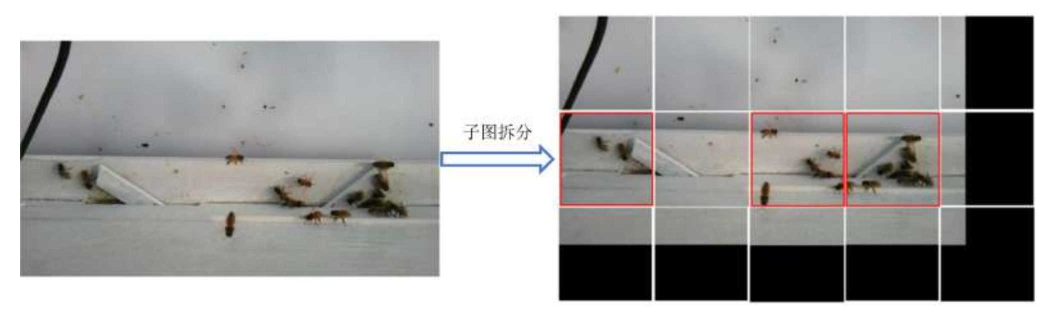


图3　尺度自适应小目标识别算法数据集生成过程示意图

Fig. 3 Generation process of new dataset of scale adaptive small objects detection method

示。若目标完全处于一个子图中，将其分配给该子图；若目标跨越相邻的两个子图（图4中标注的蜜蜂A和蜜蜂B），计算两者中目标面积较小部分的占比，若其小于设定的阈值，且将较小部分丢弃，只保留面积较大的部分（蜜蜂B）；若大于设定的阈值，则两者都保留，将其分配给各自所在的子图（蜜蜂A）并重新计算其坐标；若目标被划分为四部分，同样依据其占比确定其是否保留，并同时计算其在子图中的新坐标。原目标重分配及坐标重计算过程算法流程如图5所示。

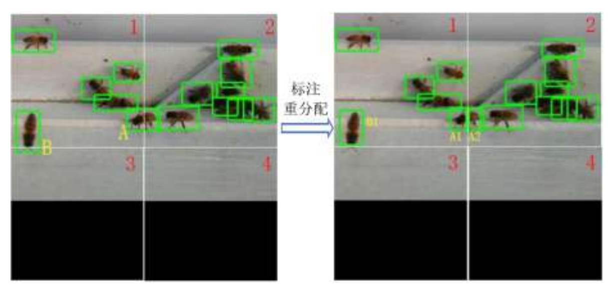


图4　尺度自适应小目标识别算法目标重分配过程示意图

Fig. 4 Process of objects reallocation of scale adaptive small objects detection method

2.3 模型的训练和使用

由于新数据集和原数据集只是在像素尺度上

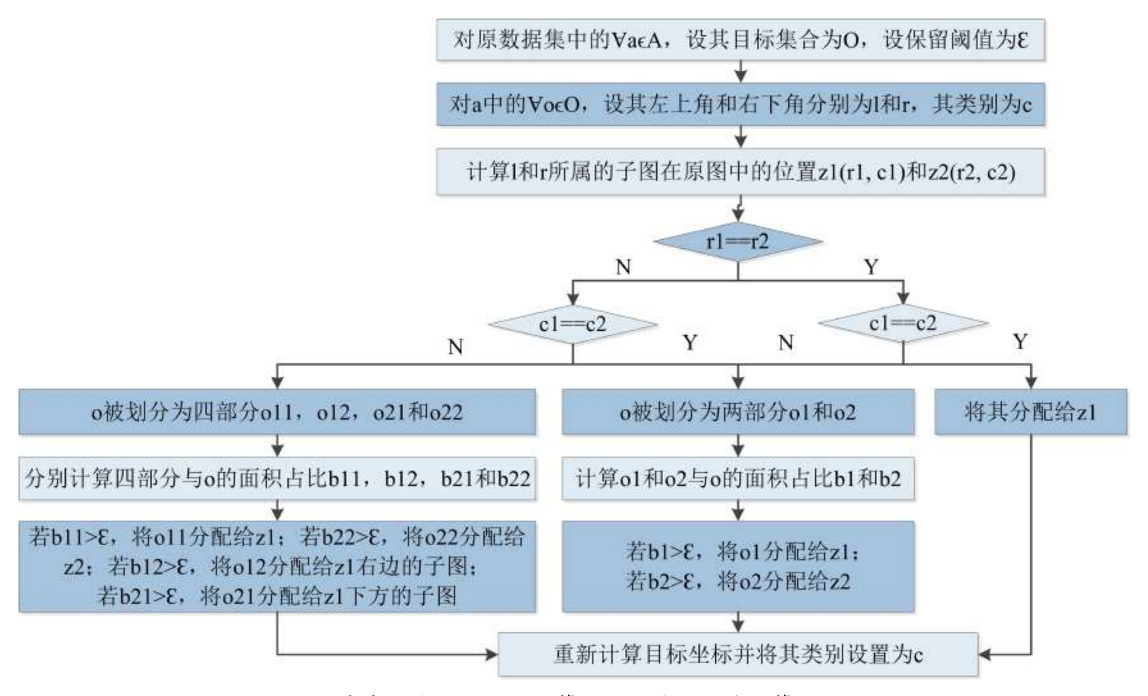


图5　尺度自适应小目标识别算法目标重分配过程算法流程图

Fig. 5 Objects reallocation algorithm flow chart for scale adaptive small objects detection method

进行了调整，目标的特征及背景不变，采用原数据集训练的识别模型已经学习到了很多目标特征，与像素尺度调整后的目标特征具有高度相似性。因此，采用迁移学习继续训练在原尺度图像训练得到的结果模型，加快模型收敛，减少模型训练的时间。

模型使用的整体流程如图6所示。由于新模型针对目标尺度较大的图像，在使用模型时同样需要将原图拆分为多个子图。为使子图衔接处的目标能被准确识别，子图之间设置一定的重叠率，重叠率的比例和目标的像素尺度相关，重叠尺度和目标尺度相似即可，过大的像素尺度会造成拆分子图数量过多，降低算法时效性。



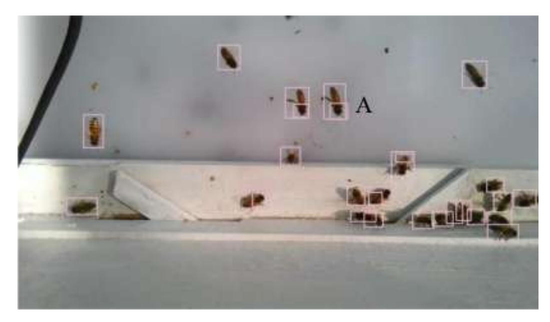
图6 尺度自适应小目标识别算法流程图

Fig. 6 Flow chart of scale adaptive small object detection algorithm

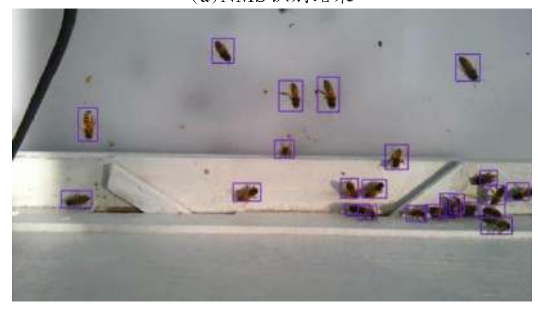
将每个子图输入新模型得到该子图中的目标集合，然后依据目标坐标及其所属子图在原图中的位置还原出目标在原图中的坐标，收集所有子图检测得到的目标集合，采用NMS移除由于模型本身产生的冗余框（图7（a））。由于相邻子图的重叠区域目标重复，可能造成识别结果中同一个目标存在内外两个识别框的问题（图7（a）中标注A）。这是因为原有的NMS采用交并比IOU（图8）作为边界框的定位精度（公式（1）），当两个边界框面积相差较大且两者相交部分和较小的边界框占比较大时，交并比的值小于设定阈值。原有的NMS不能去除此种冗余框，为了去除嵌套处不完整目标识别冗余框，采用交小比IOS（公式（2））作为两个识别框的相似性度量，研究了一种交小比非极大抑制（IOS-NMS）方法实现对其内部冗余框的去除，通过IOS-NMS后的目标识别结果见图7（b）。

$$\mathrm{IOU}=\frac{A\cap B}{A\cup B} \tag{1}$$

$$\mathrm{IOS}=\frac{A\cap B}{\min(A,B)} \tag{2}$$



(a)NMS识别结果



(b)改进的NMS二次识别结果

图7 基于NMS及IOS-NMS的目标识别结果

Fig. 7 Detections after non-maximum suppression and intersection over small non-maximum suppression

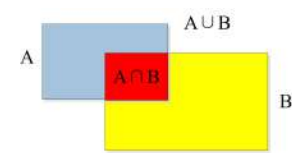


图8 两个目标边界框及它们的交集和并集

Fig. 8 Two object bounding boxes and their intersection and union

3 算法性能评估

3.1 试验设计

为验证本算法性能，以蜂场中的蜜蜂识别为例进行验证试验。试验硬件环境采用CPU型号为Intel Corei76700k，搭载一台GeForce GTX Titan X的GPU，系统操作系统为Ubuntu，采用Pytorch深度学习框架搭建网络。

对采集的2613张图像进行手工标注，创建蜜蜂图像原始数据集，图像中蜜蜂尺度均值（即蜜蜂像素数与图像总像素数的比值）为0.0037。采用2.2节中提出的拆分方法建立新的数据集，子图尺寸设置为360×320像素，共得到6269张

图像，蜜蜂尺度均值约为0.028。

选用SSD模型和YOLOv3深度学习网络模型作为算法中的目标检测模型，首先采用原始数据集训练模型，分别得到针对原始图像的原模型，然后使用迁移学习的方法使用新数据集继续训练原模型，即新尺度SSD模型和新尺度YOLOv3模型，获取针对新数据集的新模型，即尺度自适应新模型。采用同一批验证数据集进行验证分析。为避免其他因素的影响，测试中新算法的目标检测模型部分和原尺度的相应模型设置相同的置信度阈值。

3.2 性能验证结果分析

为分析子图尺寸、子图重叠率对结果的影响，分别选取300×300、500×500和700×700像素三种子图尺寸和0.2、0.05两种子图重叠率进行验证试验。采用精度、召回率和单张图像的计算时间三个指标评估模型性能，结果如表1所示。

表1　三种识别模型性能对比分析结果

Table 1 Comparison results for the three detection models

性能指标	召回率/%	精度/%	平均单张图像识别时间/s
原尺度SSD模型	94.6	87.3	0.046
原尺度YOLOv3模型	96.2	88.1	0.059
尺度自适应新模型（$zr = zc = 300$，$os = 0.2$）	98.4	89.9	0.970
尺度自适应新模型（$zr = zc = 300$，$os = 0.05$）	98.4	88.3	0.810
尺度自适应新模型（$zr = zc = 500$，$os = 0.2$）	98.4	89.9	0.512
尺度自适应新模型（$zr = zc = 500$，$os = 0.05$）	98.2	89.2	0.362
尺度自适应新模型（$zr = zc = 700$，$os = 0.2$）	97.8	89.6	0.315
尺度自适应新模型（$zr = zc = 700$，$os = 0.05$）	97.1	89.5	0.227

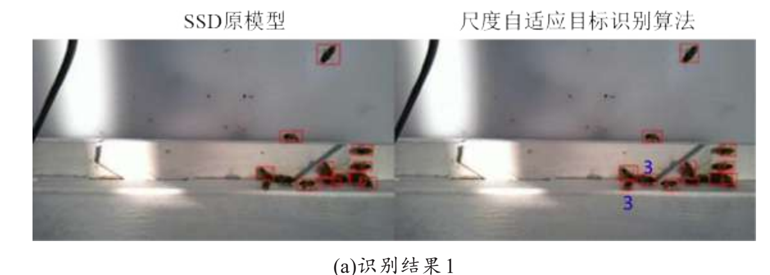
注：zr和zc分别表示子图的高度和宽度；os表示子图之间的重叠比例

由验证试验结果可知，和原尺度SSD模型相比，尺度自适应目标识别算法的召回率普遍有所提升，当拆分像素尺度为300×300和500×500，且重叠率为0.2时，召回率达到了同样的最高值98.4%，较原尺度SSD模型高3.8%。

部分目标检测结果对比图如图9所示。可知，在识别召回率方面，尺度自适应目标识别算法对特征不明显及不完整蜜蜂也能识别出来。如图9中标注1的只有局部的蜜蜂，标注2的是由于光照或者蜜蜂移动的原因造成的不清晰的蜜蜂，标注3的是由于蜜蜂的姿势及所处位置造成的蜜蜂像素尺度更小的蜜蜂。这是因为本算法对原图进行了拆分，增加了目标物的尺度，从而丰富了目标特征，能识别出不易识别的蜜蜂。当子图尺度从300×300像素增加至700×700像素时，召回率逐渐减小。总体上，召回率随着子图尺度的增大而减小。子图尺度为500×500和700×700像素、重叠率取值0.05时的召回率均低于取值0.2时的召回率，召回率随着重叠尺度的减小而减小。

在识别精度方面，尺度自适应目标识别算法平均精度普遍较原尺度SSD模型有所提高，尺度分别为300×300和500×500像素、重叠率均为0.2时，模型的精度均达到最大的89.9%，较原尺度模型的87.3%提高了2.6%。当子图尺度为300×300像素、尺度为0.05时，精度最低为88.3%。

尺度自适应目标识别算法的召回率和平均精度普遍高于原尺度YOLOv3模型。YOLOv3在小目标识别方面具有较优越的性能，使用残差网络增加了网络的深度，采用多层特征融合的方法丰富低层小目标的语义特征，当模型的输入尺度设为其默认尺度416×416像素时，其召回率为95.1%。虽然YOLOv3专门针对小目标的识别对



(a)识别结果1

(b)识别结果2

(c)识别结果3

图9 尺度自适应小目标识别算法和SSD原模型测试识别结果对比图

Fig. 9 Comparison results between the original single shot multibox detector model and the scale adaptive small object detection algorithm

网络结构进行了调整，但当小目标有效像素较少时，仍会由于特征表达不充分造成困难小目标识别不能取得满意的效果。

在识别速度方面，尺度自适应目标识别算法计算速度较原模型普遍降低，尺度从300×300增加到700×700像素时，单张图像的计算时间成倍增加。相同尺度下，重叠尺度0.2时的计算时间较0.05大约多1/5。

3.3 复杂环境下算法性能测试

为评估算法的尺度自适应性和泛化性，进一步验证本算法在复杂环境下性能，从网上爬取了三张不同尺度、不同背景下的野外蜜蜂图像进行了对比测试。图像详细信息如表2所示。

采用尺度自适应目标识别算法和SSD模型识别蜜蜂。尺度自适应目标识别算法分别采用两种子图尺度为300×300和500×500像素，重叠率均选0.2。识别结果如图10所示。

由图10可知，SSD模型在图像尺寸较小的pic1时能识别出1～2个蜜蜂。随着图像尺寸的增加，在pic2上能识别出3～8个蜜蜂，而在尺寸更大的pic3上均没能识别出一只蜜蜂。而尺度自适应小目标识别算法由于通过拆分为子图调整了蜜蜂的尺度，能更好地适应不同尺度的输入，尤其是当子图尺寸为300×300像素时，识别性能并没有因为原图尺寸的变化而有明显的下降。

由于模型训练中没有蜂巢及整体蜂箱等相关背景的图像，没能学习到相关背景的特征，同时选取的测试图像中均具有高密度的蜜蜂，所以尺度自适应目标识别算法在识别的准确率和召回率方面结果并不令人满意。但是，和SSD模型测试对比结果充分说明了尺度自适应目标识别算法具有较强的尺度适应能力和泛化性能。

表2 尺度自适应小目标识别算法测试图像信息

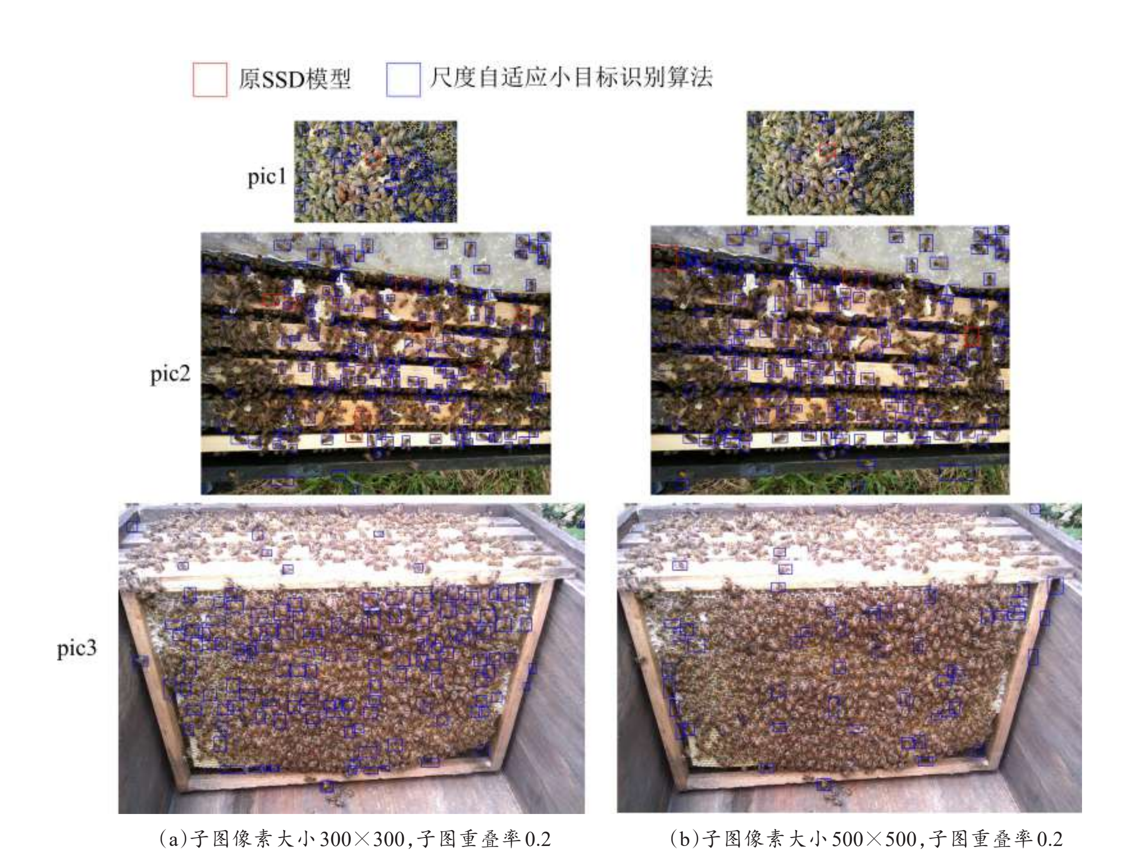
Table 2 Information for test images of scale adaptive small objects detection method

图像名称	分辨率	蜜蜂尺度均值	环境
pic1	670×420	$4.975e^{-3}$	野外蜂巢
pic2	1440×1080	$0.900e^{-3}$	野外蜂箱侧面
pic3	1920×1280	$0.559e^{-3}$	野外蜂箱全景

4 讨论与结论

4.1 讨论

尺度自适应目标识别算法通过将原图拆分为多个子图作为深度学习网络的输入，增加了目标的尺度，能提取丰富的目标特征，从而提高识别性能，尤其在目标的绝对像素数量充分且目标尺度较小时，尺度自适应目标识别算法能有效地避免输入图像归一化导致的目标有效像素数过度减小导致的目标特征提取困难的问题，能更充分体



（a）子图像素大小300×300，子图重叠率0.2　（b）子图像素大小500×500，子图重叠率0.2

图10　野外不同尺度蜜蜂图像的识别结果对比

Fig. 10 Bee detection results on images with different size in farmlands

现新算法的性能。

深度学习网络推理过程是目标识别中耗时占比最大的部分，通过拆分原图多次推理会导致算法的时效性下降，导致单张图像识别时间成倍增加，时效性随子图尺寸的减小而增加，较高的重叠率也会导致子图数量的增加造成识别时间增加。依据目标的像素数和模型的输入尺寸，选取合适的子图尺寸和重叠率能增加精度和召回率，同时提高模型的时效性。

终端采集设备多变，导致采集的图像分辨率和质量不一，尤其是农业生产中，从业人员多样，采集设备良莠不齐，若对所有尺寸的图像都一次性输入模型中，必将因为目标尺度过小且不一导致模型性能下降。本研究算法首先依据采集目标的尺度进行拆分处理，实现对所有尺度图像的自动处理，提高模型的尺度适应性和泛化能力。

4.2 结论

本研究针对农田中小目标识别困难，目标尺度多变造成的识别性能差的问题，研究了一种提高目标有效像素数量及其尺度的方法以提升模型性能。首先将原图拆分为多个子图，将每个子图作为目标检测模型的输入，然后采用二阶段非极大值抑制方法实现最终目标的计算。试验结果表明，该方法能有效识别一些特征不明显的困难目标，尺度自适应目标识别算法的召回率和精度都普遍高于原算法，召回率最高提高3.8%，精度最高提高2.6%。本算法的召回率和精度也普遍高于YOLOv3模型。但由于本算法时效性较差，适用于对召回率和精度要求较高的非实时性计算。

参考文献：

[1] NAUATA N, HU H, ZHOU G T, et al. Structured label inference for visual understanding[J]. IEEE Transac-

tions on Pattern Analysis and Machine Intelligence, 2019, 42(5): 1257-1271.
[2] 黄凯奇, 任伟强, 谭铁牛. 图像物体分类与检测算法综述[J]. 计算机学报, 2014, 36(12):1-18.
HUANG K, REN W, TAN T. A review on image object classification and detection[J]. Chinese Journal of Computers, 2014, 36(12):1-18.
[3] ZOU Z, SHI Z, GUO Y, et al. Object detection in 20 years: A survey[J/OL]. arXiv: 1905.05055v2 [cs. CV], 2019.
[4] 李科岑, 王晓强, 林浩, 等. 深度学习中的单阶段小目标检测方法综述[J]. 计算机科学与探索, 2022, 16(1): 41-58.
LI K, WANG X, LIN H, et al. Survey of one stage small object detection methods in deep learning[J]. Journal of Frontiers of Computer Science and Technology, 2022, 16(1): 41-58.
[5] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J/OL]. arXiv: 2004.10934, 2020.
[6] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J/OL]. arXiv:1804.02767 [cs.CV], 2018.
[7] MAHTO P, GARG P, SETH P, et al. Refining YOLOv4 for vehicle detection[J]. International Journal of Advanced Research in Engineering and Technology, 2020, 11(5): 409-419.
[8] ZHAI S, SHANG D, WANG S, et al. DF-SSD: An improved SSD object detection algorithm based on denseNet and feature fusion[J]. IEEE Access, 2020: 24344-24357.
[9] 奚琦, 张正道, 彭力. 基于改进密集网络与二次回归的小目标检测算法[J]. 计算机工程, 2021, 47(4): 241-247, 255.
XI Q, ZHANG Z, PENG L. Small object detection algorithm based on improved dense network and quadratic regression[J]. Computer Engineering, 2021, 47(4): 241-247, 255.
[10] SHENZ Q, LIU Z, LI J G, et al. DSOD: Learning deeply supervised object detectors from scratch[C]// The 2017 IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Computer Society, 2017: 1937-1945.
[11] 李航, 朱明. 基于深度卷积神经网络的小目标检测算法[J]. 计算机工程与科学, 2020, 42(4): 649-657.
LI H, ZHU M. A small object detection algorithm based on deep convolutional neural network[J]. Computer Engineering & Science, 2020, 42(4): 649-657.
[12] 周慧, 严凤龙, 褚娜, 等. 一种改进复杂场景下小目标检测模型的方法[J/OL]. 计算机工程与应用: 1-8. [2021-10-04]. http://kns. cnki. net/kcms/detail/11.2127. TP.20210419.1404.049.html.
ZHOU H, YAN F, ZHU N, et al. An approach to improve the detection model for small object in complex scenes[J/OL]. Computer Engineering and Applications: 1-8. [2021-10-04]. http://kns. cnki. net/kcms/detail/11.2127.TP.20210419.1404.049.html.
[13] ESTER M, KRIEGEL H P, SANDER J, et al. A density-based algorithm for discovering clusters in large spatial databases with noise[C]// The Second International Conference on Knowledge Discovery and Data Mining. Portland, Oregon, USA: AAAI, 1996: 226-231.
[14] 李云红, 张轩, 李传真, 等. 融合DBSCAN的改进YOLOv3目标检测算法[J/OL]. 计算机工程与应用: 1-12. [2021-10-04]. http://kns. cnki. net/kcms/detail/11.2127.TP.20210327.1437.002.html.
LI Y, ZHANG X, LI C, et al. Improved YOLOv3 target detection algorithm combined with DBSCAN[J/OL]. Computer Engineering and Applications: 1-12. [2021-10-04]. http://kns. cnki. net/kcms/detail/11.2127. TP.20210327.1437.002.html.
[15] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression[C]// IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, New York, USA: IEEE, 2019: 658-666.
[16] YANG Y, LIAO Y, CHENG L, et al. Remote sensing image aircraft target detection based on GIoU-YOLOv3[C]// 2021 6th International Conference on Intelligent Computing and Signal Processing. Piscataway, New York, USA: IEEE, 2021: 474-478.
[17] ZHENG Z, ZHAO H, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C]// The 34th AAAI Conference on Artificial Intelligence, the 32nd Innovative Applications of Artificial Intelligence Conference, the 10th AAAI Symposium on Educational Advances in Artificial Intelligence. Piscataway, New York, USA: AAAI, 2020: 12993-13000.
[18] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, (99): 2999-3007.
[19] 张炳力, 秦浩然, 江尚, 等. 基于RetinaNet及优化损失函数的夜间车辆检测方法[J]. 汽车工程, 2021, 43 (8): 1195-1202.
ZHANG B, QIN H, JIANG S, et al. A method of vehicle detection at night based on RetinaNet and optimized loss functions[J]. Automotive Engineering,

2021, 43(8): 1195-1202.
[20] 郑秋梅, 王璐璐, 王风华. 基于改进卷积神经网络的交通场景小目标检测[J]. 计算机工程, 2020, 46(6): 26-33.
ZHENG Q, WANG L, WANG F. Small object detection in traffic scene based on improved convolutional neural network[J]. Computer Engineering, 2020, 46(6): 26-33.
[21] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
[22] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]// In European Conference on Computer Vision. Cham, Switzerland: Springer: 2016.
[23] Neubeck A, Gool L J V. Efficient non-maximum suppression[C]// International Conference on Pattern Recognition. Piscataway, New York, USA: IEEE Computer Society, 2006: 848-855.
[24] 李景琳, 姜晶菲, 窦勇, 等. 基于Soft-NMS的候选框去冗余加速器设计[J]. 计算机工程与科学, 2021, 43 (4): 586-593.
LI J, JIANG J, DOU Y, et al. A redundacy-reduced candidate box accelerator based on soft-non-maximum suppression[J]. Computer Engineering & Science, 2021, 43(4): 586-593.
[25] 张长伦, 张翠文, 王恒友, 等. 基于注意力机制的NMS在目标检测中的研究[J]. 电子测量技术, 2021, 44(19): 82-88.
ZHANG C, ZHANG C, WANG H, et al. Research on non-maximum suppression based on attention mechanism in object detection[J]. Electronic Measurement Technology, 2021, 44(19): 82-88.

Scale Adaptive Small Objects Detection Method in Complex Agricultural Environment: Taking Bees as Research Object

GUO Xiuming, ZHU Yeping, LI Shijuan, ZHANG Jie, LYU Chunyang, LIU Shengping*

(*Agricultural Information Institute*, *Chinese Academy of Agricultural Sciences*/*Key Laboratory of Agri-information Service Technology*, *Ministry of Agriculture and Rural Affairs*, *Beijing* 100081, *China*)

Abstract: Objects in farmlands often have characteristic of small volume and high density with variable light and complex background, and the available object detection models could not get satisfactory recognition results. Taking bees as research objects, a method that could overcome the influence from the complex backgrounds, the difficulty in small object feature extraction was proposed, and a detection algorithm was created for small objects irrelevant to image size. Firstly, the original image was split into some smaller sub-images to increase the object scale, and the marked objects were assigned to the sub-images to produce a new dataset. Then, the model was trained again using transfer learning to get a new object detection model. A certain overlap rate was set between two adjacent sub-images in order to restore the objects. The objects from each sub-image was collected and then non-maximum suppression (NMS) was performed to delete the redundant detection boxes caused by the network, an improved NMS named intersection over small NMS (IOS-NMS) was then proposed to delete the redundant boxes caused by the overlap between adjacent sub-images. Validation tests were performed when sub-image size was set was 300×300, 500×500 and 700×700, the overlap rate was set as 0.2 and 0.05 respectively, and the results showed that when using single shot multibox detector (SSD) as the object detection model, the recall rate and precision was generally higher than that of SSD with the maximum difference 3.8% and 2.6%, respectively. In order to further verify the algorithm in small target recognition with complex background, three bee images with different scales and different scenarios were obtained from internet and test experiments were conducted using the new proposed algorithm and SSD. The results showed that the proposed algorithm could improve the performance of target detection and had strong scale adaptability and generalization. Besides, the new algorithm required multiple forward reasoning for a single image, so it was not time-efficient and was not suitable for edge calculation.

Key words: object detection; machine vision; small object; farmland; bee; single shot multibox detector; YOLOv3